Ceilings or Floors? :

Gender Wage Gaps by Education in Spain[*]

Sara de la Rica*, Juan J. Dolado** & Vanesa Llorens***

(*) Universidad del País Vasco & IZA
(**) Universidad Carlos III & CEPR & IZA
(***) LECG Consulting Spain S.L.

Forthcoming in *Journal of Population Economics*.

ABSTRACT

This paper analyses the gender gap throughout the wage distribution in Spain using data from the ECHP. Quantile regression and panel data techniques are used to estimate wage equations at relevant percentiles in a given representative year (1999), and over time (1994-2001). In contrast with the steep increasing pattern found in other countries, the flatter evolution of the Spanish gender gap hides an intriguing composition effect when the sample of workers is split by education. For high-educated workers, in line with the conventional *glass ceiling* hypothesis, the gap increases as we move up the distribution. However, for less-educated workers the gap decreases. This declining pattern is even more acute when we correct for selection and remains similarly shaped when differences in characteristics are accounted for. We label this novel fact as a *floor* pattern and argue that it can be explained by statistical discrimination exerted by employers in countries where less-educated women have low participation rates. Such a hypothesis is further confirmed when the panel structure of the ECHP is exploited.

JEL Classification: J16 and J71.

Keywords: gender gap, glass ceilings, floor pattern, quantile regressions, panel data

[*] Corresponding author: Juan J. Dolado (dolado@eco.uc3m.es). We would like to thank three anonymous referees, an Editor, M. Arellano, S. Bentolila, F. Felgueroso, J. Gardeazábal, M. Jansen, B. Petrongolo, seminar participants at Amsterdam, Marseille, País Vasco (Bilbao), Toulouse, CEMFI, CES, ECARES, ESEM 2004 (Madrid) and SAE 2004 (Pamplona). Special thanks go to C. García-Peñalosa for her insightful comments and suggestions. The first two authors gratefully acknowledge financial support from the Spanish Ministry of Education (SEC2003-04826; SEC2004-04101) and the EC (MRTN-CT-2003-50496).

1. Introduction

It is a widely documented fact that men earn higher wages than women (*gender gap*), even after controlling both for observable characteristics related to their productivity and the overall wage structure (see, e.g., Blau and Kahn, 2000). There is an extensive literature about this topic based on the comparison of average (log) wages. Nonetheless, the analysis of the gender gap at other points of the wage distribution has drawn less attention.[1] Lately, however, there has been a growing interest in examining whether the gap increases throughout the distribution, in line with the so-called *glass ceiling* hypothesis.

In this paper, inspired by the approach in Albrecht et al. (2003) to document the existence of glass ceilings in Sweden, we first derive quantile measures of the gender gap in Spain at the end of the 1990s to next check our preferred explanation for its evolution using panel data. This is an interesting issue, since Spain, like some other Southern-Mediterranean countries (e.g., Greece or Italy but not Portugal), still has a much lower female participation than the Nordic countries and therefore patterns of women's achievements in the labour market are bound to differ markedly from those found there.[2] Indeed, the evidence we provide here supports this view: the overall gender gap is much less steep in Spain than in Sweden. Moreover, our most important contribution is to uncover an interesting a composition effect behind this pattern when the sample of workers is split by education.[3] As we move up the wage distribution, the gender gap for workers with high (college/tertiary) education (henceforth, H-group) *increases*. On the contrary, for workers with less (primary/secondary) education (L-group), the gap *decreases*. Whereas the increasing pattern for the H-group mimics the previous findings by Albrecht et al (2003) for Sweden, the negative slope of the gap for the L-group, to the best of our knowledge, is a novel fact in the literature with interesting policy

---

[1] Well-known examples of this type of research are Chamberlain (1994) and Buchinsky (1994, 1998a, b) who use quantile regressions to analyze the overall wage structure in the U.S. An application of this method to Spain can be found in Abadie (1997). Applications to gender wage discrimination are Fitzenberger et al. (2001), Newell and Reilly (2001) and Albrecht et al. (2003). As regards the latter topic in Spain, there are two related studies to ours. On the one hand, García et al. (2001), using the 1991 Encuesta de Conciencia, Biografía y Estructura de Clase, control both for the endogeneity of education and the selection of women into the labor market and conclude tthat the discrimination component is higher at the top than at the bottom of the wage distribution. On the other, Gardeazábal and Ugidos (2005), making use of the 1995 Encuesta de Salarios, also find that the raw gender gap increases along the distribution but, by contrast, estimate that the discrimination component, in relative terms, is larger at the bottom of the distribution.

[2] The Spanish female activity rate (% of population aged 15-64) in 2001 was 50.7% whereas it reached 75.7 % in Sweden and 60.2% in the EU. By educational levels, the corresponding rates in Spain were 80.4% and 48.0% for the women with tertiary education and less than tertiary education (84.6% and 68.3% in Sweden), respectively (see OECD, 2002). Indeed, the group of Spanish working women is formed by very heterogeneous cohorts. Since the 1980s, female participation has surged (from 33.3% in 1980 until 50.7% in 2001) mainly due to an increase in access to higher education and a reduction in fertility rates (see, e.g., Arellano and Bover, 1995).

[3] Educational attainments are treated as predetermined categories throughout the paper since our goal is not to estimate its returns.

implications. In the sequel, in order to stress its differences with the conventional glass ceiling pattern, we will use the catch word *floor* pattern to coin this new phenomenon.[4]

A preliminary illustration of these facts is provided by means of the 1999 (6th. wave) of the European Community Household Panel (ECHP, henceforth) for full-time workers.[5] Figure 1a plots the gender gap (in terms of the differences of (logged) gross hourly wages of male and female workers) in Spain throughout the distribution, together with the mean gap (dashed line).[6] There is a decreasing trend that becomes stable around the 60th percentile to then increase sharply at the higher quantiles. As expected, the gap at the mean differs notably from the gap at the various percentiles. This non-monotonic evolution contrasts with the one found for Sweden (see Figure 1b) where the raw gap increases by 35 percentage points from the bottom to the top of the distribution (see Albrecht et al. , 2003).

[Figures 1a, 1b, 1c, 1d about here]

Figures 1c and 1d, in turn, depict the corresponding quantile gender gaps for the Spanish L and H-groups. The graph for the H-group fits well with the conventional glass ceiling shape. In stark contrast to this shape, the gap for the L-type workers is decreasing. Aggregation of both types of workers therefore leads to the flatter and non-monotonic pattern depicted in Figure 1a. Hence, there seems to be an intriguing *educational composition* effect that deserves greater scrutiny. Interestingly, northern and central European countries, such as Denmark and the U.K. (Figures 2a and 2b), exhibit monotonically increasing gaps for both groups, as in Sweden, while in southern European countries with lower female participation, like Greece and Italy (Figures 2c and 2d), the behavior is more irregular. Nonetheless, the gaps for the L-group in both countries resemble the *floor* pattern found for Spain.[7]

[Figures 2a, 2b, 2c, 2d about here]

---

[4] Indeed, a more appropriate name would be "glass ceilings at the ground floor" since it refers to gender pay gap at the bottom quantiles of the wage distribution. Since a "glass floor" could be wrongly interpreted as preventing women´ s wages from falling too low, the term floor patterns will be used in the sequel for the sake of brevity. Notice that it should not be confused with the "sticky floors" concept related to the lower wages that women receive at the top of the distribution due to the lack of alternative job offers relative to men (see Booth et al., 2003). In a non-competitive context where rents are shared between workers and firms, a higher male reservation wage could raise their wages above those of equally productive women.

[5] Similar patterns hold for the other ECHP waves.

[6] The compared percentiles correspond to the wage distributions of men and women separately. If we were to consider the position of women in the men's distribution, it is found that 31% (5.4%) of women are in the bottom (top) decile.

[7] The reported gender gaps for Denmark, U.K, Greece and Italy also correspond to the 1999 wave of the ECHP. The Swedish gender gap is reproduced from Figure 1 of Albrecht et al. (2003) which corresponds to 1998 with the data coming from *Statistics Sweden* (SSW). Activity rates by education in those countries can be found in Table A1 in the Appendix.

Several explanations arise in order to reconcile these divergent patterns by educational attainments: [8]

1. L-type women´s activity rates are still much lower in southern Mediterranean countries than in northern and central European countries, despite the increase in participation that has taken place in the former countries during the last two decades (see Table A1 in the Appendix). Low participation rates can lead to a non-random selection of women into work seriously affecting the resulting gender gap. Therefore, the *floor* pattern could be just the consequence of a pure selection phenomenon. One could think of two alternative ways in which selectivity biases could operate.[9] On the one hand, if more experienced women - typically placed, as we will argue below, at the middle and top of the distribution- are a positively selected group, then the floor pattern could be explained by this sort of selectivity. On the other, we may think of negative selection of women as experience and tenure increase. This would be the case if most women work at the beginning of their careers when they are young, but only those with higher economic needs remain in the labour market. This type of selection would exacerbate the negative slope of the gap and cannot explain the *floor* pattern observed in the data. As will be argued in section 3, our evidence supports the latter effect.

Hence, a plausible explanation for the *floor* pattern should rely on arguments different from selection. The one we propose relies upon statistical discrimination when specific training is a requirement to perform a job. The idea is that insofar as less-educated women's careers in these countries suffer from frequent interruptions - due to societal discrimination in family duties, lack of family-aid policies, or religious beliefs - employers may use statistical discrimination in wage-setting in order to pay a lower proportion of the training cost for women than for men. This phenomenon results in lower women´ s wages at the lower part of the distribution which typically captures wages paid at entry jobs. However, as their job tenure expands, the reasons behind statistical discrimination vanish and women´s wages will converge to men's, conditional on equal productivity.

2. *H-type women*, in contrast, have much higher participation rates - only slightly below men´ s even in Southern Mediterranean countries (see Table A1) - and are less likely to quit given the large human capital investment that they have undertaken.

---

[8] Another possible explanation, not mentioned below, could arise from some form of unobserved heterogeneity affecting L-type women in relation to their male counterparts (see more about this in section 5). Further, the OECD (2002) warns about the possibility of having measurement errors in the survey stemming from the fact that the interviewed persons provide direct information about their own wages, rather than their employers, as is the case with matched employer-employee data. If those earning more, mainly men, have a larger propensity to understate their wages, the gap for the higher quantiles would be underestimated. Although this argument could imply a downward bias of the gap at the top of the distribution for both groups of workers, it cannot explain the pattern found at the bottom of the distribution for the L-group.

[9] In either case, selectivity bias affects the mean gender gap in Southern-Mediterranean countries; cf. Olivetti and Petrongolo, 2005. The important issue is, however, how it affects the slope of the gap throughout the wage distribution.

Therefore, neither selection nor our proposed explanation for the *floor* pattern is bound to apply in this case. In line with their presumed higher commitment, their wages will be similar to men's at their entry jobs. As we move up, however, women's wages may fall below men's if the traditional glass ceiling phenomenon holds, for which several explanations have been proposed in the literature. [10]

To the extent that the glass ceiling phenomenon is well documented elsewhere, we will focus in the rest of the paper on what constitutes a novel fact in this literature, namely the *floor* pattern for countries with low participation rates of L-type women. Spanish data is used to evaluate our interpretation.[11]

Some preliminary evidence on the plausibility of our proposed explanation can be drawn from a recent survey carried out by the Spanish Ministry of Labour (MTAS) on the performance of working women in Spain (Instituto de la Mujer, 2005). In particular, the report yields information about the relationship between voluntary quits and educational attainments across genders. Based on a stratified sample of 4.000 individuals aged 16-65 from the Spanish Labour Force (EPA) in 2003, it is reported that the proportion of quits (of at least one year) among L-type women is 31.7% whereas the corresponding rate for L-type men is 14.4%. By contrast, the quit rates for H-type workers are 13.3% and 8% respectively. Thus, the low participation rate for L-type women relative to men ´s seems to be positively correlated with their propensity to quit. This correlation seemingly supports employers´ beliefs about women exiting employment faster than men and reinforces statistical discrimination as an equilibrium phenomenon.

Since the gender gap displayed in Figure 1d could be attributed to a *lower productivity* of women or to a *lower market return* for given characteristics, it is important to disentangle these two components. To do so, we follow two alternative econometric approaches: one using quantile regressions (QR) - controlling for selection – in cross-sectional data for a given representative year (1999, for comparison with Albrecht et al. ´s (2003) results, and another using panel data based on the eight available waves (1994-2001) of the ECHP to control for unobserved heterogeneity.[12] With both

---

[10] For example, Polachek (1981) predicts that women choose occupations where the cost of career interruptions is low. The existence of occupational segregation by gender would support this argument; cf. Dolado et al. (2004). Another explanation relies upon the fact that women have a lower probability to be promoted to jobs with higher responsibilities even in the case where they have both the same ability distribution than men. The model by Lazear and Rosen (1990) confers a higher productivity in the household to women, an assumption that makes employers reluctant to invest in their training on an equal basis with men. Hence, only the more productive women would be promoted. Finally, as mentioned earlier, the interpretation of the "sticky floors" model by Booth et al. (2003) relies upon men receiving a larger number of alternative offers.

[11] Details of the H-group sample and empirical analysis on the glass ceiling phenomenon using QR methods can be found in an earlier working-paper version of this work; cf. de la Rica et al. (2005).

[12] Given the cross-sectional nature of the data in the first approach, the interpretation of results in terms of our proposed explanation requires the underlying assumption - for which we provide some supporting

approaches we find that longer female tenure is disproportionately rewarded vis-á-vis men´s in the L-group, consistent with our statistical discrimination hypothesis whereby, as job tenure of L-type women expands, the gap in favour of equally productive men gets eroded.

The rest of the paper is organized as follows. In Section 2, we motivate our analysis by offering a simple theoretical model that is consistent with our statistical discrimination explanation of the *floor* pattern pertaining to the L-group. Section 3 is devoted to describing the QR econometric methodology, the data employed, the effects of selectivity corrections and the results of the gender regressions. In Section 4 we perform the QR gender gap decomposition. Section 5 exploits the panel dimension of the ECHP to provide further support to our statistical discrimination hypothesis. Finally, Section 5 concludes. An Appendix offers a detailed description of the data.

2. A model of statistical discrimination for the floor pattern

To account for the presence of *floors* in the quantile evolution of the gender gap for the L-group, we use a simple model motivated by Acemoglu and Pischke (1998) ´s analysis of the financing of training in frictional labour markets, adapted to our framework where no search frictions are present but where exogenous disutility shocks are allowed to induce quits.

Let us assume that L-type workers of both genders live for two periods and are endowed with the same ability which, for simplicity, is normalized to 1. It is also assumed that workers need to get specific training to perform a job. Thus, the need to consider two periods: period 1, where workers get trained and period 2, where workers produce. The amount of training, τ, is decided by the firms which hire the workers.

In the initial period, workers receive the training and do not produce. However, they receive an initial wage, W_1, to cover workers´ living expenses. Provision of training implies that firms bear a linear investment cost, $c(\tau)= \tau$, and that workers produce a quantity of output in period 2 according equal to $a(\tau)$, which is a concave function given by $a(\tau)= \beta\, \tau^{\alpha/2}$ with $0<\alpha<1$, so that $a'(.)>0$ and $a''(.)<0$. In the second period, workers receive a disutility shock, ω, which may force them to quit the job (say, for family duties). The ω shock is an *i.i.d.* random variable with c.d.f. $F(\omega)$ which is revealed to the worker after the wage in the second period, W_2, has been set by the firm. Thus, workers will always get trained in period 1 and will produce in period 2 as long as $W_2 - \omega \geq 0$. Further, we assume that there is free entry of firms in the market.

The key difference between men and women is that the c.d.f. for men, $F_m(\omega)$, is stochastically dominated by the c.d.f. for women $F_f(\omega)$, namely $F_m(\omega) > F_f(\omega)$ for $\omega > 0$. This assumption captures the fact that women are more likely to be affected by the shock than men, perhaps because they have higher outside opportunities at home production or

evidence in section 3 - that age/experience and tenure increase as we move up the distribution, so that lower quantiles are likely to reflect wages at initial stages of workers´ careers while higher quantiles correspond to wages at later stages. However, this assumption is not needed in the second approach where we can follow individuals in their jobs over time.

by societal discrimination. To simplify the algebra, and without loss of generality in terms of the qualitative results, we will assume that dF(.) are uniform distributions, such that the density functions verify: $f_m(\omega) = U[0, \varepsilon_m]$ and $F_f(\omega) = U[0, \varepsilon_f]$, with $\varepsilon_f > \varepsilon_m$.

To solve for both wages and the amount of training, we proceed backwards in time. Under the assumptions that the wage in period 2, W_{2i} (i=f, m), is offered before ω is realized, firms will choose W_{2i} in order to maximize expected profits in period 2, $\Pi_{2i}(\tau_i)$, namely

$$\max_{W_{2i}} \int_0^{W_{2i}} \left(a(\tau_i) - W_{2i}\right) dF_i(\omega) = \max_{W_{2i}} \frac{1}{\varepsilon_i}\left[a(\tau_i)W_{2i} - W_{2i}^2\right], \quad i = f, m, \tag{1}$$

whereby the first-order condition (f.o.c.) w.r.t. W_{2i} implies that the wage paid in equilibrium to male and female workers are $W_{2m}^* = a(\tau^*_m)/2$ and $W_{2f}^* = = a(\tau^*_f)/2$, respectively, where τ^*_i, *i=m, f* are the optimal amounts of training for each gender also chosen by the firm in period 1.[13] By replacing W_{2i}^* in the bracketed term in (1), notice that the firm ′s profits in period 2 (Π^*_{2i}) when hiring men and women are given by $\Pi^*_{2m} = a^2(\tau^*_m)/4\varepsilon_m$ and $\Pi^*_{2f} = a^2(\tau^*_f)/4\varepsilon_f$.

Denoting by (Π_{12i}) the sum of profits in both periods, i.e., $\Pi^*_{12i} = \Pi^*_{2i} - \tau^*_i - W^*_{1i} = \dfrac{a^2(\tau^*_i)}{12\varepsilon_i} - \tau_i^* - W_{1i}^*$, the f.o.c. of Π^*_{12i} w.r.t. to τ, implies that τ^*_i satisfies the condition $a^2(\tau^*_i)/12\varepsilon_i = a(\tau^*_i)/2a'(\tau^*_i)$. Using the functional form $a(\tau)=\beta\tau^{\alpha/2}$, it is straightforward to obtain that $\tau^*_i = (\alpha\beta^2/4\varepsilon_i)^{1/(1-\alpha)}$ and $W_{2i}^* = (\beta/2)(\alpha\beta^2/4\varepsilon_i)^{\alpha/2(1-\alpha)}$. Notice that $\tau^*_f < \tau^*_m$ if and only if $\varepsilon_f > \varepsilon_m$. Thus, women receive less training than men given their higher quit probability. Then, by defining the (logged) gender wage gap in period 2 as $GW_2 = (\ln W_{2m}^* - \ln W_{2f}^*)$, it follows that GW_2 is positive and equals $\alpha(\ln\varepsilon_f - \ln\varepsilon_m)/2(1-\alpha) > 0$.

Next, having chosen W_{2i}^* and τ^*_i, free entry of firms in the market implies that firms choose the wages in period 1, W_{1i}^*, so as to equate overall profits in both periods (Π_{12i}) to zero, that is

$$\Pi^*_{12i} = \Pi^*_{2i} - \tau^*_i - W^*_{1i} = \frac{a^2(\tau^*_i)}{12\varepsilon_i} - \tau_i^* - W_{1i}^* = 0, \tag{2}$$

which yields,

[13] This is just the average of the worker s productivity′ and the outside wage which is assumed to be zero. The weight ½ in the average is due to the choice of the uniform distribution in the illustration. Alternative distributions will give rise to a weighted average with unequal weights.

$$W_{1i}^{*} = \frac{1-\alpha}{\alpha}\tau^{*}{}_{i} = \frac{1-\alpha}{\alpha}\left(\frac{\alpha\beta^{2}}{4\varepsilon_{i}}\right)^{1/1-\alpha} \tag{3}$$

Given the higher female quit probability, the (logged) gender wage gap in period 1, GW_1 = ($\ln W_{1m}$* - $\ln W_{1f}$*) is also positive and equals $(\ln\varepsilon_f - \ln\varepsilon_m)/(1-\alpha)$>0.[14] Finally, comparing GW_1 and GW_2, the following result holds.

Proposition: *The gender wage gap in period 1 is larger than in period 2, i.e., GW_1-GW_2=[(2-α)/2(1- α)] $(\ln\varepsilon_f - \ln\varepsilon_m)$>0.*

The intuition for this proposition is quite simple. Since the disutility shock is not known at the time when W_{2i} is offered, the best that firms can do is to match this outside offer by setting a wage equal to a fraction of the observed productivity $a(\tau^{*}{}_{i})$ which, under a uniform distribution, equals $a(\tau^{*}{}_{i})$/2. Hence, firms will obtain a surplus equal to Π^*_{2i}=$a(\tau^{*}{}_{i})$ - W_2*= $a(\tau^{*}{}_{i})$/2 in period 2. Since expected profits in period 2 are higher with the more-stable male workers than with the less- stable female workers (because $a(\tau^{*}{}_{m}) > a(\tau^{*}{}_{f})$) firms pay a higher wage to men than to women in the initial period, despite having the same productivity. This explains (i) why women receive a lower wage in period 1 than men, and (ii) why the gender gap in period 1 is larger than in period 2.

Which are the empirical implications of these results? Given that entry jobs for L-type workers typically have wages located at the lower part of the distribution, the previous results suggest that for this type of workers the gender gap is higher at the bottom than at the top of the distribution, leading to the *floor* pattern. Moreover, at entry jobs (period 1 in the model), L-type women are statistically discriminated against L-type men. However, as women remain in the firm (i.e., they have longer tenure), employers´ statistical discrimination of women decreases. Thus, they offset their former discriminatory behavior by increasing stable-women´s rewards relative to men´s, for equal productive characteristics. [15]

---

[14] Notice that for H-type workers $\varepsilon_m = \varepsilon_f$ and, hence, the wage gap in periods 1 and 2 is zero. Thus, the glass ceiling pattern for this type of workers has to be explained by some other theories like those discussed in footnote 10 that we do not deal with in this paper.

[15] This explanation somewhat mimics the standard one available in the literature about statistical discrimination, concerning the employers´ private learning process about workers´ ability. As the employer learns more about the worker through a longer tenure in the job, the return on education (the signal) decreases while the return on experience/tenure increases (see, e.g., Farber and Gibbons, 1996).

3. Quantile regressions: Methodology, Data and Results

3.1 Methodology (QR)

Following Koenker and Bassett (1978) and Buchinsky (1998a), the model of QR in a (log) wage-equation setting can be described as follows. Let (w_i, x_i) be a random sample, where w_i denotes the (log) hourly gross wage of individual i, x_i is a vector K ×1 of regressors, and $Q_\theta(w_i|x_i)$ is the conditional θ^{th} quantile of the distribution of w_i given x_i. Then, under the assumption of a linear specification, the model can be defined as

$$w_i = x_i'\beta_\theta + u_{\theta i}\ , \qquad Q_\theta(w_i|x_i) = x_i'\beta_\theta \tag{4}$$

where the distribution of the error term $u_{\theta i}$, $F_{u\theta}(\cdot)$, is left unspecified, just assuming that $u_{\theta i}$ satisfies $Q_\theta(u_{\theta i}|x_i) = 0$. The estimated vector of QR coefficients, $\hat{\beta}_\theta$, is interpreted as the marginal change in the conditional quantile θ due to a marginal change in the corresponding element of the vector of coefficients on x, and can be obtained using the optimization techniques described in Koenker and Bassett (1982).

Buchinsky (1998b) has extended model (4) to account for selectivity bias. As is conventional in standard parametric selectivity correction at the mean, two different wages ought to be considered: (i) a reservation wage, w^R_i, which depends linearly on a vector of characteristics z_i , and (ii) a wage offer, w^*_i , which also depends linearly on a vector x_i , such that z_i contains at leat a continuous variable that is not included in x_i. Since the wage offer is only observed for the individuals for whom $w^*_i > w^R_i$, we have that $w_i = Dw^*_i$ where D is the usual indicator function $I(w^*_i > w^R_i)$. Under mild conditions discussed in Buchinsky (1998b), the probability of working $P(w^*_i > w^R_i /z_i)$ is a function of a known index g_i ($=z_i'\gamma$), and the observed wage equation becomes

$$w_i = x_i'\beta_\theta + h_\theta(g_i) + \varepsilon_{\theta i}\ , \quad Q_\theta(\varepsilon_{\theta i} /x_i, D=1) = 0, \tag{5}$$

where $h_\theta(g_i)$ is approximated by a power series expansion of the inverse of the Mill´s ratio. To make estimation of (5) feasible, the unknown coefficients γ are replaced by their estimates obtained from a first-stage minimization of the squared distance between D_i and a (non-parametric) kernel of the conditional expectation $E(D_i /z_i, \gamma)$. [16]

3.2 Data and Results

The data are drawn from the 1999 (6th. wave) of the ECHP which provides information in a harmonized format for the EU countries on earnings, employment, and many other socio-demographic variables. The information is obtained from surveys to a fixed panel of households (70,000 in the EU and around 7,000 in Spain) since 1994. Our sample is restricted to full-time employed L-type workers aged 16-64, excluding self-employed, full-time students and those in the military service, leading to 1,585 men and 726 women. Appendix A contains a detailed description of the variables used in the regressions while

[16] In the estimation reported in section 3.2, the non-parametric kernel is a truncated normal and two terms are used in the power series expansion.

Table 1 offers summary descriptive statistics of the sample.[17] Furthermore, descriptive statistics for non-participating women in the L-group are also reported in the last two columns of this Table since they will be used to correct for non-random selection. As can be observed, L-type male workers are much more experienced than women (4 years) - where the precise definition of experience is given below- they have a longer tenure (1.8 years) and are older (1.7 years). Women have a larger share in firms with less than 20 employees and work more often in the public sector. The mean gender gap is around 23%. As regards non-working women, they are older (2.2 years) than working women, a larger fraction has young children and their non-labour household income is larger. This last feature may point out to higher economic needs of working women.

The first two rows in Table 2 present the evolution of experience and tenure throughout the main quantiles, confirming that they both increase monotonically as we move up the distribution. When comparing the 10th and 90th quantiles, experience and tenure increase by 32 and 20 years, respectively. Thus, to motivate the results in this section, our underlying assumption about the link between the quantiles of wages and stages in a worker´ s job career, seems plausible for this dataset. The third row also reports the fraction of temporary contracts across quantiles. It shows that, despite a declining pattern, the difference in tenure for workers above and below the median cannot be solely explained by the former group holding a permanent contract and the latter a temporary one.[18].

[Tables 1 and 2 about here]

We have estimated QR equations (reported at the 10th, 25th, 50th, 75th and 90th quantiles) where the (log. of) gross hourly wage is regressed on different subsets of covariates. Heteroskedastic-robust estimation at the conditional mean has also been undertaken for comparative purposes. As is conventional in *mincerian* wage equations, the controls in each of the two educational groups are: experience and its square, experience interacted with age of children, tenure in the current job, marital status, age of children and secondary education. At this stage, it is important to notice that actual work experience is not available in the ECHP. Instead, the best we can do with the available data is to use a measure of potential experience computed as the current age of an individual minus the age at which he/she started his/her first job which is a better proxy

[17] Descriptive statistics for non-working men in the L-group are not reported since their participation rate is high (see Table A1). Moreover, attempts to correct for self-selection proved inconsequential: the selectivity term was highly insignificant and the remaining coefficients on observable characteristics hardly changed relative to the uncorrected regressions.

[18] We are grateful to a referee for raising the issue of whether the relationship between low tenure and holding a temporary contract could affect the validity of our statistical discrimination explanation. We think that this is not the case for at least four reasons. First, because we control for type of contract in the wage regressions. Secondly, because an interaction term between tenure and type of contract never proved to be significant. Thirdly because, since 1997, temporary contracts with a maximum fixed-duration of 3 years were abolished; c.f. Dolado et al. (2002). And, finally, as will be discussed below, because the counterfactual decompositions performed in section 4 indicates than temporary contracts play a much smaller role than tenure in explaining the gender gap.

for true women´s experience than the standard age minus years of school minus 6. To further improve the measurement of true experience, following Buchinsky (1998b), our definition of potential experience is interacted with a dummy of dependent children aged below 16, assuming that the main alternative use of working time is child rearing. To account for the demand side of the labour market, regional dummies and size of municipality have also been included. Further controls are firm size, immigrant condition, type of contract (permanent or temporary), sector (private or public), supervisory role and 15 occupational dummies. The latter, also used in Albrecht et al. (2003), are arguably endogenous, yet they are useful in explaining the gender gap from an ”accounting exercise” viewpoint, as these authors point out .[19]

As a preliminary check, we started the analysis by running a pooled OLS regression at the mean and pooled quantile regressions at the above-mentioned quantiles for the joint sample of male and female workers. A (female) gender dummy captures the extent to which the gap remains unexplained after controlling for individual differences in the observed characteristics with returns restricted to be the same for both genders. Although the results of this pooled estimation are not shown to save space, it is important to report that the intercept for the gender dummy turned out to be always negative and significant, decreasing in absolute value as we move up the distribution. However, these results are only tentative since the null hypothesis of equal coefficients on the covariates for both genders is rejected with a p-value of 0.006. Thus, separate estimations for each gender are needed.

Next, in order to control for non-random selection of workers, we estimate separate QR with the Buchinsky correction described in section 3.1. Given that non-random selection is only significantly different from zero for women, we do not report the results for the participation equation for men. The results of estimating this equation for L-type women are presented in Table 3a. The dependent variable equals 1 if working and 0 otherwise and the independent variables are age, number of children, age of children, age of parents living at home, marital status, being immigrant and total household non-labour income. For comparative purposes, the first column reports the coefficient from a probit estimation while the second column reports the estimates from the single-index model.[20] Results indicate that older women tend to participate less, as is the case of having larger household non-labour income, having old parents at home or being married. By contrast, having completed a secondary degree or being immigrant increases the probability of working.

[Table 3a about here]

Tables 3b (men) and 3c (women) report the separate QR. We first present selection-corrected QRs for women, and uncorrected QRs for men. Further, to check the effects of

[19] Unfortunately, the ECHP does not provide information on parents’ education or occupation, which could provide appropriate instruments to correct for endogeneity.

[20] As explained in Buchinsky (1998b, pg.7), the constant and the coefficient on one of the continuous variables (e.g., age) are not identified in a single-index model. Hence, they are normalized by setting them equal to their values in a probit model so that the results are comparable.

selectivity, estimates of uncorrected wage equation for women are also shown in Table 3d. Starting with the comparison between men and women, Tables 3b and 3c reveal that the coefficient on age/experience is larger for men than for women and that the gap decreases as we move up the distribution. Further, having a secondary educational attainment yields a higher return for women at the lower quantiles, as is also the case of working in the public sector or having a permanent contract. The coefficients on being married are larger for men, particularly at the lower quantiles. The most interesting finding, however, is that the estimated coefficients on tenure are higher for women than for men at the lower quantiles but converge to the same value as we move up in the distribution. Note that this pattern can be easily interpreted in terms of our proposed explanation based on statistical discrimination. In effect, employers perceive that the job attachment of L-type women is lower than men´s, particularly at the early stage of their careers which, according to Table 2, is likely to be captured by the bottom of the wage distribution. Thus, the reward to longer tenure should be higher for them than for the "more stable" men.

Comparison of the results in Tables 3c and 3d allows us to check the effects of correcting for non-random selection in the sample of women. The first effect to notice is that both the constant term and the coefficients on several dummies in Table 3c fall relatively to those reported in Table 3d , implying an estimated average gap of 22.6% instead of the 19.7% obtained without correction. Furthermore, this reduction is much larger at the bottom of the distribution than at the top. Finally, the declining pattern of the estimated coefficient on tenure has a much more pronounced negative slope in the corrected equation than in the uncorrected one. Thus, overall, correcting for selectivity makes the *floor* phenomenon even more acute. This result rules out the alternative explanation of the *floor* pattern based on a pure selection phenomenon, whereby women with wages in the middle and upper parts of the distribution are a favourably selected sample in terms of characteristics. On the contrary, it supports that low-wage women are the favourably selected sample, exacerbating in this fashion the negative slope of the gender gap throughout the wage distribution. Moreover, the different patterns of the estimated coefficients on tenure by gender yield some support to our statistical discrimination hypothesis.

[Tables 3b, 3c and 3d about here]

Summing up, the evidence presented so far points out that: (i) returns to observable characteristics differ by gender, (i) these differences change as we move throughout the distribution, and (iii) selectivity correction exacerbates the *floor* pattern. The next step is to investigate how important are unobservables in explaining the gender gap in order to check the consistency of our statistical discrimination hypothesis.

4. Decomposition of the gender gaps

4.1 Methodology (MM decomposition)

A generalization of the Oaxaca-Blinder decomposition to a QR framework has been proposed by Machado and Mata (MM)'s (2005) using Monte Carlo methods. The decomposition is based on the construction of a counterfactual distribution of w^f which represents the distribution of female wages that would have prevailed if women had been endowed with their own characteristics but were paid like men.[21] This counterfactual distribution is denoted $F(\tilde{w}^f / x^f, \hat{\beta}_\theta^m)$, where $\tilde{w}^f$ are generated values of w^f and $\hat{\beta}_\theta^m$ are the (male) quantile regression coefficients.

The steps in the MM algorithm in order to construct $F(\tilde{w}^f / x^f, \hat{\beta}_\theta^m)$ can be summarized as follows:

- For each quantile θ=0.01, 0.02,.., 0.99, estimate the quantile regression vector of coefficients $\hat{\beta}_\theta^m$ using the male dataset.
- Use the female dataset to generate fitted values, $\tilde{w}^f(\theta) = \hat{\beta}_\theta^{m\prime} x^f$. For each θ, this generates N^f fitted values, where N^f is the size of the female sample. Next, randomly select s=100 of the elements of $\tilde{w}^f(\theta)$ for each θ and stack these into a 99x100 element vector, $\tilde{w}^f$. The empirical c.d.f. of these values is the estimated counterfactual distribution, namely what women would have earned if they were paid like men.
- Then compare the counterfactual distribution with the empirical male and female wage distributions whose θ^{th} quantiles are defined by $w^m(\theta)$ and $w^f(\theta)$, respectively. The gender gap at the θ^{th} quantile can be decomposed as: $w^m(\theta)$ - $w^f(\theta) = [\, w^m(\theta) - \tilde{w}^f(\theta)\,] + [\, \tilde{w}^f(\theta) - w^f(\theta)\,]$. The first term in brackets is the *characteristics effect* since it measures the contribution of different covariates to the gender gap at the θ^{th} quantile. The second term in brackets is the *returns effect* since it captures the contribution of differences in returns to the gender gap at the θ^{th} quantile.[22]

By randomly re-sampling the male data 250 times, using the bootstrap method by Parzen et al. (1994), bootstrapped standard errors for the contribution of these components have been obtained.

A similar decomposition procedure is applied to the wage residuals obtained after correcting for tenure in the QR. The idea behind this second counterfactual gap is to check whether the *floor* pattern becomes less intense when the effect of tenure is omitted. If the pattern becomes less steep, it must be that tenure is one of the main driving factors

---

[21] Differences in observed characteristics are typically evaluated in the decomposition at the men´ s returns, under the assumption that they are not distorted by discriminatory behaviour.

[22] Notice that by implementing this decomposition, in contrast to Albrecht et al. (2003), we are evaluating the difference in characteristics at the market returns of men. By interchanging the role of men and women in the MM procedure, which is what these authors do, we can obtain the alternative evaluation at women´ s rewards, $\hat{\beta}_\theta^f$, so that the *Returns* component is evaluated at the male dataset for each quantile. The results of this alternative decomposition are not presented here but the qualitative findings about the unexplained gap remain similar.

behind the *floor* phenomenon. By contrast, if it remains similar, tenure can not be an important explanation. Additionally, in order to compare the role of tenure with that of temporary/permanent contracts, we perform a similar exercise, this time correcting for type of contract in the QR.

4.2 Results of the MM decomposition

Table 4 presents the results of the MM decomposition with all covariates, while Figure 3 depicts the raw and the three generated counterfactual gaps (solid line with all covariates and two dashed lines with wage residuals). Although implementing this decomposition with the selectivity corrected wage equations only adds slight additional computational burden to the MM procedure (since the single-index estimation only has to be done once; see Albrecht et al, 2005), we only report results just for the uncorrected equations. We do so for two reasons. First, as discussed in section 3, because the *floor* pattern persists after controlling for selection. And, secondly, because one could argue that the selection correction part of the MM decomposition is not really necessary, since the estimated counterfactual distribution uses men ´s quantile regression coefficients and these are likely not biased by selection,

As can be observed, the counterfactual gap is clearly decreasing along the distribution, reaching a minimum of about 50% at the 75th quantile, in accord with the *floor* phenomenon. That is, while the raw gap is basically explained by differences in returns at the bottom of the distribution, differences in observed characteristics explain about one-half of the gap at the top of the distribution. Interestingly, the counterfactual gap with the wage residuals after correcting only for tenure is much flatter. By contrast, the counterfactual gap correcting only for type of contract looks fairly similar to the one obtained with all covariates. Thus, in accord with our proposed explanation, statistical discrimination seems to be an important factor in driving the large gap at the bottom of the distribution for which we provide further support in the next section using an alternative econometric approach.

[Table 4 and Figure 3 about here]

5. Testing the statistical discrimination hypothesis with panel data

In the previous sections we have used QR to argue that statistical discrimination provides a plausible explanation for the *floor* pattern in the L-group under the assumption that age and tenure jointly increase as we move up the wage distribution holds. The evidence provided in Table 2 supports this interpretation. However, our conjecture could be greatly reinforced if we were to follow individuals in their jobs *over time.*

Fortunately, this can be done by exploiting the panel dimension of the ECHP in its eight available waves (1994-2001). Panel data estimation (in particular, fixed effects estimation) of similar wage equations to the ones estimated above - this time in real terms- can be used to test the significance and sign of the interaction of tenure and gender (female dummy) in each educational group, and for young and older workers, respectively. Since fixed-effects estimation removes all individual unobserved heterogeneity which does not change over time, it allows us to analyze in a more direct way the effect of longer tenure on wages of workers with different educational attainments.

Given that our statistical discrimination hypothesis for the *floor* pattern, we should observe greater early-career returns to tenure for L-type women relative to L-type men. These differences, however, should not apply to the H-type group where both men´ s and women ´s participation rates are high. Our empirical strategy relies upon separate and pooled estimation of men´ s and women´ s wage equations. In the pooled estimation, all explanatory variables are interacted with a female dummy. Then, the coefficient on the interaction term *Tenure* Female* for the L-group should be positive and statistically significant whereas it should not be significant for the H-group. Moreover, the return to an extra year of tenure of a L-type woman is bound to increase much more when they are young than when they are older since young women are more likely to be involved in child-rearing tasks which raise their probability of quitting. Thus, as a further robustness check to our results, we also carry out the fixed-effects estimation splitting the sample of workers into two age groups: below 36 and above 40 years of age at the first wave[23]. We should expect a much larger coefficient on the previous interaction term for the former age group than for the latter. Finally, to control for selectivity biases in the sample of L-type workers, we implement a two-step Heckman-correction procedure whereby a first-stage participation probit (not reported) is estimated separately for men and women using similar identifying covariates to those discussed in Table 3c.

Table 5a presents the descriptive statistics regarding (log) wages, age and tenure of the overall sample. Only workers that are observed working at least twice are considered in order to estimate the effect of changes in tenure on wages. The final sample consists of an unbalanced panel that contains 21657 observations, out of which 11678 belong to the L-group and 17535 to the group below 36 years of age as of 1994. For the group of young workers (first panel of Table 5a), we can observe that the average gender gap is higher in the L-group than in the H-group (0.17 versus 0.07), age is around 0.5 years higher for men than for women in both groups, and the difference in tenure is larger for L-type workers (on average, men have around 0.5 years longer tenure than women) than for H-type workers (on average, men have around 0.33 years longer tenure than their female counterparts).

The results from the fixed-effects estimation of the wage equations for younger and older workers are reported in Tables 5b and 5c, respectively. For each age group and for each educational level we present separate estimations by gender plus a pooled estimation where all explanatory variables are interacted with a female dummy. We report the

[23] Small changes in the age boundaries of each of the groups do not make a significant difference in the results.

effects of age (and its square) and tenure (and its square) for men and for women separately and their interactions with a female dummy in the pooled estimation. [24] As shown in Table 5b, the returns to age are lower for women than for men in both educational groups. Regarding tenure, however, we can observe- both from the separate gender wage regressions and from the pooled estimation of men and women- that the returns to tenure for the L-type women are significantly higher than for men. The estimated coefficients on the linear *Tenure*female* term and its square are highly significant and the differential effect is concave. On the contrary, the corresponding estimated coefficients in the H-group are non significant. Thus, both sets of results heavily support the role of statistical discrimination in explaining the *floor* phenomenon. This is further confirmed by the results in Table 5c for older workers where the differential gender effect on the returns to tenure in either educational group is statistically insignificant.

To evaluate the differential effect on tenure between the younger women and men in the L group, we can use the estimated coefficients on the interaction term *Tenure*Female* and its square in the penultimate column of Table 5b. At an average tenure of around 3.2 years (see Table 5a), the difference is 3.46 percentage points per year, quite a significant gap. More importantly, the cumulative effect of the difference in tenure and tenure-squared coefficients is maximized at 9 years, at which point this cumulative difference amounts at 8.1 percentage points.

Finally, another result worth mentioning is that the probability of working exerts a positive effect on wages only for the young L-type women, meaning that these women that we observe working (at least twice) are not a random sample of the population of all L-type women. Rather, they have higher wages than average. Thus, as pointed out in the QR analysis, this means that the *floor* pattern would be even stronger if our sample of women were a random sample of the female population. Notice as well that, to the extent that unobserved heterogeneity is assumed to be time invariant, fixed-effects estimation weakens alternative explanations of the *floor* pattern based on heterogeneity affecting women (relatively to men) in a disproportionate way.

[Tables 5a,5b and 5c about here]

---

[24] Age replaces potential experience as a control because the age at which the individual started the first job does not change over time.

6. Conclusions and policy implications

In this paper, we have analyzed the evolution of gender gaps along the wage distribution in Spain both using QR and panel data approaches. Our main finding is that behind an irregular evolution of the gap for the whole sample of individuals, there is distinctive difference between the patterns of the gaps when distinguishing by high and low educational attainments. While for the H-group the gender gap is increasing over the distribution (*glass ceiling*), as in many other countries, it is strongly decreasing for the L-group (*floor pattern*). In particular, this last pattern -not documented before- contrasts with those found in central and northern European countries, where the gap increases as we move up the distribution irrespectively of educational attainments. Moreover, we present descriptive evidence arguing that a similar pattern also holds in other southern European countries, like Greece or Italy, where female labour market participation of L-type women is still rather low.

We argue that non-random selection can not account for the novel *floor* phenomenon. Instead we favour an alternative explanation based on statistical discrimination in the presence of training costs co-financed by employers and workers. Due to the historical low participation of women in the L-group, employers may use statistical discrimination to lower their wages vis-à-vis more stable men in the lower part of the wage distribution since they expect future career interruptions to jeopardize their financing of specific training. However, as their job tenure expands the reasons behind statistical discrimination vanish, so that their wages converge to men's wages with the same characteristics. This statistical discrimination hypothesis is further reinforced when we exploit the panel dimension of the ECHP, following the gender gaps as women stay longer in the same firms.

The policy implications stemming from the existence of *floors* could be far reaching. For example, one could think that they give rise to a multiple equilibrium model of gender gaps in wages and participation rates. If cultural beliefs or reduced social expenditure on family aid lead to low L-type female participation, our explanation predicts that this initial situation will increase the wage gap in favour of men which, in turn, will exacerbate lower female participation. Hence, both effects lead to a "bad" equilibrium with low participation and large gaps. On the contrary, if societal discrimination is absent or family aid is generous, high female participation will lead to lower wage gaps which will further feed back into larger female participation driving the economy to a "good" equilibrium. This last situation seems to correspond to the Nordic countries while the former applies to some of the above-mentioned southern Mediterranean countries. In these circumstances, public policies in favour of reconciling family and work could move an economy from the "bad" to the "good" equilibrium not only, as is conventionally thought, by increasing female participation but, as stressed in this paper, by reducing the gender wage gap as well. Further analysis of these policy implications when participation is endogenously determined is part of our current research agenda (see, de la Rica et al., 2006).

Appendix

A.1: Definition of variables

The variables are drawn from the 1999 (6th wave) of the ECHP. Our group of interest is composed by wage earners working full-time (at least 1560 hours in a year or 30 hours per week on average). In this section we provide a more detailed description of the variables used in the analysis.

Gross hourly wage: The ECHP collects data on average monthly labor income (gross and net), from salaried workers. Labor income includes salary bonus (divided by working months), and overtime. When a worker has more than one job, only the main job income is considered. Weekly hours in the main job are available, including overtime hours. We have set an upper bound of 60 hours to this variable in order to minimize the self-declared bias. This correction affects 2% of men and 0.9% of women from our total sample. Then, gross hourly wage is the monthly gross salary divided by 52/12 and multiplied by the weekly hours worked in the main job.

Experience: defined as current age minus age at which the individual started working.

Exp*Children: interaction between experience and a binary variable that takes a value of 1 when an individual has dependent children (from 0 to 16 years). In the participation equation in Table 3a, we considered separately the cases in which children are between 0 and 11 years (Exp*Children 0-11) or between 12 and 16 years (Exp*Children 12-16), but only the first one showed up significant.

Secondary education: dummy for having completed upper secondary education**.**

Individual characteristics: dummies for marital status, immigrant condition, district of residence and district size.

Type of contract: temporary or permanent.

Sector: private or public.

Supervisory role: directive or managing position, supervisor of at least another employee and without responsibility for the rest of employees.

Tenure: obtained as the difference between the year of the survey, 1999, and the year of the start of the current job.

Firm size: from 1 to 4 employees, from 5 to 19 employees, from 20 to 49 employees, from 50 to 99 employees, from 100 to 499 employees and above 500 employees.

Occupation: fifteen occupational groups have been considered, corresponding to an intermediate level of aggregation of the ISCO-88 (COM) classification. The list is: Legislators, senior officials and managers (OC1); Physical, mathematical, engineering, life science and health professionals (OC2); Teaching professionals (OC3); Other professionals (OC4); Physical, mathematical, engineering, life science and health associate professionals (OC5); Teaching and other associate professionals (OC6); Clerks (OC7), Models, salespersons and demonstrators (OC8); Personal and protective services workers (OC9); Skilled agricultural and fishery workers (OC10); Extraction and building trades workers, other craft and related trades workers (OC11); Metal, machinery, precision, handicraft printing and related trades workers (OC12); Plant and machinery operators and assemblers (OC13); Sales and services elementary occupations (OC14); and Agricultural, fishery and related laborers, laborers in mining, construction, manufacturing and transport (OC15).

Table A1: Labour activity rates by educational attainment (2002)

Countries	Men			Women		
	Less than secondary	*Secondary*	*Tertiary*	*Less than secondary*	*Secondary*	*Tertiary*
Denmark	75.4	87.3	92.7	55.8	80.7	88.4
Sweden	78.0	87.9	90.4	65.1	83.4	88.1
U.K.	65.9	88.1	92.2	50.7	76.4	87.3
Greece	81.0	88.5	89.7	42.1	57.2	82.4
Italy	75.5	86.1	90.9	34.8	67.9	82.7
Spain	83.5	90.1	91.9	42.3	67.6	83.1

Source: OECD, Employment Outlook, 2002

References

1. Abadie, A. (1997), "Changes in the Spanish Labor Income Structure during the 1980´s: A Quantile Regression Approach", *Investigaciones Económicas*, 21, 253-272.

2. Albrecht, J., A. Björklund and S. Vroman (2003), "Is There a Glass Ceiling in Sweden?", *Journal of Labor Economics*, 21, 145-177.

3. Albrecht, J., A. van Vuuren, and S. Vroman (2004), "Decomposing the Gender Gap in the Netherlands with Sample Selection Adjustments", (mimeo).

4. Acemoglu, D. and J.S. Pischke (1998), "Why Do Firms Train: Theory and Evidence", *Quarterly Journal of Economics*, 113, 79-111.

5. Arellano, M. and O. Bover (1995), "Female Labour Force Participation in the 1980s: The Case of Spain", *Investigaciones Económicas*, 29, 171- 194.

6. Blau, .F. and L. Kahn (2000), "Gender Differences in Pay", *Journal of Economic Perspectives*, 14, 75-99

7. Booth, A., M. Francesconi and J. Frank (2003), "A Sticky Floors Model of Promotion, Pay and Gender", *European Economic Review*, 47, 295- 322.

8. Buchinsky, M. (1994), "Changes in the U.S. Wage Structure 1963-1987: An Application of Quantile Regression", *Econometrica*, 62, 405-458.

9. Buchinsky, M. (1998a), "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research", *The Journal of Human Resources*, 33, 88-126.

10. Buchinsky, M. (1998b), "The Dynamics of Changes in the female Wage Distribution in the USA: A Quantile Regression Approach", *Journal of Applied Econometrics*, 13, 1-30.

11. Chamberlain, G. (1994), "Quantile Regression, Censoring and the Structure of Wages", *Advances in Econometrics, Sixth World Congress* (Volume 1); C. A. Sims (ed.), Cambridge University Press.

12. de la Rica, S., Dolado, J. and V. Llorens (2005), " Ceilings and Floors: Gender Wage Gaps by Education in Spain", IZA DP no. 1483.

13. de la Rica, Dolado, J.J. and C. García-Peñalosa (2006), "Gender Wage and Participation Gaps: A Positive Theory of Family Benefits and Affirmative Action" (in progress).

14. Dolado, J.J., Garcia-Serrano, C. and J.F. Jimeno (2002, "Drawing Lessons from the Boom of Temporary Jobs in Spain", *The Economic Journal*, 112, 270-295.

15. Dolado, J.J., F. Felgueroso and J.F. Jimeno (2004), "Where do Women Work?: Analyzing Patterns in Occupational Segregation By Gender", *Annales d'Economie et de Statistique*, 71-72, 293-315.

16. Farber, H. and R. Gibbons (1996). "Learning and Wage Dynamics." *Quarterly Journal of Economics*, 111, 1007-1047.

17. Fitzberger, B., Koenker, R., and J.A.F. Machado (Eds.) (2001), *Economic Applications of Quantile Regression*, Physica-Verlag.

18. García, J., P.J. Hernández and A. López-Nicolás (2001), "How Wide is the Gap? An Investigation of Gender Wage Differences using Quantile Regression", *Empirical Economics*, 26, 149-168.

19. Gardeazábal, J. and A. Ugidos (2005), "Gender Wage Discrimination at Quantiles", *Journal of Population Economics*, 18, 165-179..

20. Instituto de la Mujer (2005), *Estudio sobre la Conciliación de la Vida Familiar y la Vida laboral en España: Situación Actual, Necesidades y Demanda*, MTAS. (http://www.mtas.es/mujer/mujeres/estud_inves/estud_anterior04.htm.)

21. Koenker, R. and G. Bassett (1978), "Regression Quantiles", *Econometrica*, 46, 33-50.

22. Koenker, R. and G. Bassett (1982), "Robust Tests for Heteroskedasticity on Regression Quantiles", *Econometrica*, 50, 43-61.

23. Lazear, E. and S. Rosen (1990), "Male-Female Wage Differentials in Job Ladders", *Journal of Labour Economics*, 8, 106-123.

24. Machado, J.A.F. and J. Mata (2005), "Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression", *Journal of Applied Econometrics,* 20, 445-465.

25. Manning, A. (2003), *Monopsony in Motion*, Princeton University Press.

26. Mincer, J. and S. Polacheck (1974), "Family Investment in Human Capital: Earnings of Women", *Journal of Political Economy*, 82, 76-108.

27. Newell, A. and B. Reilly (2001), "The Gender Pay Gap in the Transition from Communism: Some Empirical Evidence", *Economic Systems*, 25, 287-304.

28. OECD (2002), “Women at Work: Who Are They and How Are They Faring?”, *Employment Outlook*, 63-125.

29. Olivetti, C. and B. Petrongolo (2005), “Unequal Pay or Unequal Employment ?: A Cross-country analysis of Gender Gaps”. (mimeo)

30. Parzen, M, Wei, L. and Z. Ying (1994), “A Resampling Method Based on Pivotal Estimating Functions”, *Biometrika*, 81, 341-350.

31. Polacheck, S. (1981), “Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure”, *Review of Economics and Statistics*, 62, 60-69.

Table 1: Descriptive statistics. Low-educated workers. Spain (1999)

	M		Working F		Non-working F	
	Average	*St. dev*	*Average*	*St. dev*	*Average*	*St. dev*
N. obs.	1585		726		998	
Age	37.89	11.31	36.19	11.33	38.81	8.39
Children 0-11	0.30	0.46	0.21	0.41	0.53	0.50
Children 12-16	0.16	0.37	0.17	0.38	0.25	0.43
Age Groups						
17 to 24	0.13	0.33	0.18	0.38	0.14	0.35
25 to 34	0.31	0.46	0.32	0.47	0.38	0.49
35 to 44	0.25	0.43	0.23	0.42	0.43	0.50
≥ 45	0.31	0.46	0.27	0.45	0.05	0.21
Married	0.68	0.47	0.53	0.50	0.73	0.44
Immigrant	0.04	0.07	0.02	0.08	0.01	0.09
Secondary ed.	0.29	0.46	0.38	0.49	0.24	0.43
Weekly hours	42.64	6.19	40.36	5.74		
Gross Hourly wage	1037	487	851	484		
Log (wage)	6.86	0.41	6.63	0.47		
Experience	20.77	12.41	16.78	11.85		
Tenure	8.28	7.87	6.48	7.23		
Occupation						
OC1	0.02	0.14	0.01	0.11		
OC2	0.00	0.00	0.00	0.00		
OC3	0.00	0.03	0.00	0.04		
OC4	0.00	0.06	0.00	0.07		
OC5	0.02	0.13	0.02	0.14		
OC6	0.05	0.21	0.06	0.24		
OC7	0.07	0.25	0.18	0.38		
OC8	0.08	0.28	0.18	0.38		
OC9	0.04	0.20	0.13	0.33		
OC10	0.02	0.15	0.01	0.10		
OC11	0.22	0.42	0.08	0.27		
OC12	0.10	0.30	0.01	0.10		
OC13	0.20	0.40	0.06	0.24		
OC14	0.05	0.21	0.20	0.40		
OC15	0.13	0.33	0.06	0.24		
Firm Size						
1-4 employees	0.18	0.38	0.22	0.42		
5-19 employees	0.31	0.46	0.25	0.44		
20-49 employees	0.16	0.37	0.16	0.37		
50-99 employees	0.10	0.30	0.11	0.31		
100-499 employees	0.12	0.33	0.14	0.35		
> 500 employees	0.12	0.33	0.11	0.32		
Public sector	0.13	0.34	0.18	0.38		
Supervisory role						
Directive	0.05	0.23	0.03	0.16		
Supervisor	0.16	0.36	0.08	0.28		
Without responsibility	0.79	0.41	0.89	0.31		
Permanent contract	0.64	0.48	0.60	0.49		
Parents 65-75	0.05	0.21	0.06	0.24	0.05	0.22
Parents >76	0.02	0.15	0.05	0.22	0.03	0.17
Non-labour income	23944	65727	28065	193603	31534	68140

Table 2
Experience, Tenure and Temporariness throughout the Wage Distribution
(L-group, Spain, 1999)

	Average	**θ=10**	**θ=25**	**θ=50**	**θ=75**	**θ=90**
MEN (L)						
Experience	20.8	5.2	11.3	20.2	31.1	38.2
Tenure	8.3	0.5	1.2	3.4	15.0	20.3
% Temp. contracts	36.0	43.5	39.3	35.4	32.3	30.4
WOMEN (L)						
Experience	16.9	2.2	7.2	15.3	26.2	34.3
Tenure	6.5	0.2	0.8	3.0	11.2	19.8
% Temp. contracts	40.0	45.2	40.1	36.3	33.1	30.9

Table 3a:
Estimates of the probability of working
L-group. Spain. 1999

WOMEN	**Probit**	**Single-index**
Constant	-0.334***	-0.334
	(0.004)	(.)†
Age	-0.045***	-0.045
	(0.002)	(.)†
No. of children	-0.053***	-0.060**
	(0.016)	(0.024)
Children 0-11	-0.085***	-0.078***
	(0.024)	(0.031)
Parents >65	0.074**	0.098**
	(0.035)	(0.048)
Secondary	0.067***	0.074***
	(0.012)	(0.014)
Married	-0.092***	-0.083***
	(0.007)	(0.011)
Immigrant	0.034**	0.045**
	(0.016)	(0.021)
Non-lab.income/100	-0.056***	-0.046***
	(0.009)	(0.012)
No. obs.	1,724	1,724

Note:† The constant and age coefficients are normalized.

Table 3b. OLS and QR
L-group. Spain. 1999
Dependent variable : Ln. gross hourly wage

MEN	**Average**	**θ=10**	**θ=25**	**θ=50**	**θ=75**	**θ=90**
Constant	1.385***	1.157***	1.295***	1.379***	1.459***	1.557***
	(0.047)	(0.062)	(0.063)	(0.058)	(0.064)	(0.069)
Experience	0.013***	0.017***	0.010***	0.010***	0.008*	0.008
	(0.003)	(0.006)	(0.003)	(0.003)	(0.004)	(0.005)
Experience^2	-0.0002***	-0.0003**	-0.0002***	-0.0002***	-0.0001	-0.0001
	(0.00005)	(0.0001)	(0.00007)	(0.00007)	(0.00009)	(0.0001)
Exp*Children	-0.002**	-0.003*	-0.002	-0.001	-0.001	-0.001
	(0.0008)	(0.001)	(0.001)	(0.0009)	(0.0009)	(0.001)
Secondary ed.	0.060***	0.081**	0.060**	0.055***	0.025	0.023
	(0.020)	(0.040)	(0.027)	(0.019)	(0.024)	(0.035)
Immigrant	-0.143	-0.192	-0.222	-0.244	0.005	0.151
	(0.144)	(0.130)	(0.150)	(0.165)	(0.270)	(0.265)
Public sector	0.020	0.006	0.033	0.032	0.036	0.054
	(0.030)	(0.049)	(0.038)	(0.027)	(0.040)	(0.047)
Permanent contract	0.065***	0.105**	0.061**	0.028	0.037	0.073**
	(0.022)	(0.048)	(0.030)	(0.026)	(0.027)	(0.037)
Supervisory role						
Directive	0.161***	0.193***	0.113**	0.166***	0.135**	0.104*
	(0.040)	(0.073)	(0.051)	(0.050)	(0.068)	(0.059)
Supervisor	0.089***	0.083**	0.080***	0.096***	0.092***	0.074**
	(0.023)	(0.035)	(0.028)	(0.031)	(0.032)	(0.035)
Tenure	0.012***	0.015***	0.014***	0.012**	0.009*	0.010*
	(0.003)	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)
Married	0.079***	0.123***	0.083***	0.070***	0.073**	0.077**
	(0.021)	(0.038)	(0.025)	(0.027)	(0.030)	(0.033)
Nº Obs.	1585	1585	1585	1585	1585	1585
R^2	0.303	0.237	0.250	0.282	0.327	0.350

Note: ***, **, * represent significance at 99, 95 and 90% respectively. Standard deviation in parentheses. Dummy variables for region, firm and local council size and occupations are also included. Omitted group: wage earners in private sector in less-than-5-employees firms, without supervisory role, single, with primary education and in non-qualified jobs in services and commerce (OC14)

Table 3c. OLS and QR
(with correction for selectivity)
L-group. Spain. 1999

Dependent variable : Ln. gross hourly wage

WOMEN	**Average**	**θ=10**	**θ=25**	**θ=50**	**θ=75**	**θ=90**
Constant	1.267***	1.087***	1.252***	1.263***	1.438***	1.546***
	(0.0481)	(0.076)	(0.070)	(0.073)	(0.069)	(0.069)
Experience	0.006**	0.006	0.007	0.008*	0.009*	0.009*
	(0.003)	(0.007)	(0.006)	(0.005)	(0.005)	(0.006)
Experience^2	-0.0001	0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	(0.00007)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Exp*Children	-0.003***	0.0007	-0.006***	-0.004*	-0.003*	-0.002
	(0.001)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
Secondary ed.	0.104***	0.096**	0.081**	0.051*	0.026	0.025
	(0.033)	(0.050)	(0.039)	(0.030)	(0.039)	(0.044)
Immigrant	-0.512***	-0.283**	-0.387***	-0.556***	-0.597***	-0.606***
	(0.066)	(0.137)	(0.103)	(0.131)	(0.132)	(0.142)
Public sector	0.097***	0.150**	0.134**	0.066**	0.067*	0.068*
	(0.036)	(0.073)	(0.066)	(0.031)	(0.040)	(0.042)
Permanent contract	0.098***	0.194***	0.160***	0.111**	0.092*	0.120**
	(0.033)	(0.062)	(0.060)	(0.048)	(0.049)	(0.051)
Supervisory role						
Directive	-0.041	-0.007	0.013	0.083	0.085	0.062
	(0.108)	(0.197)	(0.129)	(0.131)	(0.179)	(0.173)
Supervisor	0.069*	0.075	0.073	0.086**	0.065	0.063
	(0.041)	(0.081)	(0.061)	(0.043)	(0.061)	(0.072)
Tenure	0.024***	0.033***	0.029***	0.013***	0.009	0.011
	(0.003)	(0.006)	(0.006)	(0.005)	(0.007)	(0.009)
Married	0.060**	0.064*	0.067*	0.093***	0.022	0.086
	(0.024)	(0.040)	(0.043)	(0.032)	(0.033)	(0.044)
Nº Obs.	726	726	726	726	726	726
R^2	0.3323	0.385	0.379	0.396	0.456	0.471

Note: ***, **, * represent significance at 99, 95 and 90% respectively. Standard deviation in parentheses. Dummy variables for region, firm and local council size and occupations are also included. Omitted group: wage earners in private sector in less-than-5-employees firms, without supervisory role, single, with primary education and in non-qualified jobs in services and commerce (OC14)

Table 3d. OLS and QR
(without correction for selectivity)
L-group. Spain. 1999

Dependent variable : Ln. gross hourly wage

WOMEN	**Average**	**θ=10**	**θ=25**	**θ=50**	**θ=75**	**θ=90**
Constant	1.308^{***}	1.118^{***}	1.295^{***}	1.313^{***}	1.447^{***}	1.555^{***}
	(0.051)	(0.082)	(0.074)	(0.078)	(0.074)	(0.073)
Experience	0.007^{*}	-0.0001	0.009	0.008	0.006	0.006
	(0.004)	(0.010)	(0.006)	(0.005)	(0.006)	(0.007)
Experience^2	-0.0001	0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	(0.00009)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Exp*Children	-0.002	0.00004	-0.004^{**}	-0.003^{*}	0.002	0.002
	(0.001)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
Secondary ed.	0.113^{***}	0.103^{*}	0.077^{*}	0.059^{*}	0.028^{**}	0.027
	(0.033)	(0.059)	(0.043)	(0.040)	(0.038)	(0.040)
Immigrant	-0.479^{***}	-0.249^{*}	-0.358^{***}	-0.528^{***}	-0.584^{***}	-0.668^{***}
	(0.073)	(0.140)	(0.112)	(0.148)	(0.144)	(0.173)
Public sector	0.108^{***}	0.150^{**}	0.134^{**}	0.066	0.067	0.066
	(0.039)	(0.073)	(0.066)	(0.051)	(0.060)	(0.072)
Permanent contract	0.121^{***}	0.194^{***}	0.160^{***}	0.111^{**}	0.090^{**}	0.104^{**}
	(0.035)	(0.062)	(0.060)	(0.048)	(0.045)	(0.048)
Supervisory role						
Directive	-0.050	0.018	0.017	-0.150	0.079	0.060
	(0.129)	(0.210)	(0.138)	(0.127)	(0.184)	(0.176)
Supervisor	0.075^{*}	0.081	0.094	0.096^{**}	0.070	0.067
	(0.045)	(0.088)	(0.061)	(0.049)	(0.066)	(0.078)
Tenure	0.021^{***}	0.027^{***}	0.024^{***}	0.016^{***}	0.013^{**}	0.010
	(0.004)	(0.005)	(0.006)	(0.006)	(0.007)	(0.010)
Married	0.065^{**}	0.040	0.071^{*}	0.093^{***}	0.022	0.086
	(0.027)	(0.050)	(0.043)	(0.032)	(0.033)	(0.044)
Nº Obs.	726	726	726	726	726	726
R^2	0.308	0.372	0.372	0.385	0.421	0.462

Note: ***, **, * represent significance at 99, 95 and 90% respectively. Standard deviation in parentheses. Dummy variables for region, firm and local council size and occupations are also included. Omitted group: wage earners in private sector in less-than-5-employees firms, without supervisory role, single, with primary education and in non-qualified jobs in services and commerce (OC14)

Table 4
Counterfactual gender gaps
L-group (Spain. 1999).

	Mean	**θ=10**	**θ=25**	**θ=50**	**θ=75**	**θ=90**
Observed Gap	**22.73**	**33.33**	**24.71**	**17.31**	**16.82**	**18.94**
Counterfactual gap	**17.08**	**31.34**	**21.54**	**11.18**	**8.76**	**9.28**
	(1.32)	(2.27)	(1.61)	(1.52)	(1.56)	(2.26)
%	75.1	94.0	87.2	64.6	52.1	49.0

Note: Standard deviation in parenthesis. The standard deviations have been obtained through 250 replications of the MM decomposition

Table 5a: Means and St.Deviation of Log Real Wages, Age and Tenure (in years)

	Workers younger than 36 at first wave (1994)				*Workers older than 40 at first wave (1994)*			
	Men		**Women**		**Men**		**Women**	
	Low	*High*	*Low*	*High*	*Low*	*High*	*Low*	*High*
Log Wage	1.82	2.21	1.65	2.14	1.92	2.76	1.72	2.48
	(0.41)	(0.49)	(0.43)	(0.49)	(0.38)	(0.56)	(0.45)	(0.42)
Age	28.77	30.62	28.18	30.16	50.45	49.21	49.84	47.63
	(5.88)	(4.87)	(5.62)	(4.76)	(5.83)	(5.65)	(5.77)	(4.50)
Tenure	3.51	4.42	3.03	4.08	4.65	7.78	5.36	8.19
	(4.33)	(4.45)	(3.86)	(4.31)	(4.88)	(5.46)	(4.91)	(5.63)
N. obs.	7933	2909	3805	2888	2266	438	1174	244

Notes: *Low*: Highest educational level achieved is secondary education or less; *High*: Highest educational level achieved is tertiary education.

Table 5b: Log Real Wage Regressions – Workers younger than 36 at first wave (1994) Fixed Effects Estimation with Selection Correction

	Men			Women			Pooled Men and Women	
	All	*Low*	*High*	*All*	*Low*	*High*	*Low*	*High*
Age	0.173	0.169	0.185	0.134	0.139	0.127	0.169	0.186
	(0.008)	(0.009)	(0.024)	(0.011)	(0.014)	(0.022)	(0.009)	(0.023)
Age^2	-0.002	-0.002	-0.002	-0.001	-0.002	-0.001	-0.002	-0.002
	(0.0001)	(0.0001)	(0.0004)	(0.0002	(0.0002)	(0.0004)	(0.0001)	(0.0003)
Tenure	0.009	0.004	0.019	0.017	0.018	0.016	0.004	0.019
	(0.002)	(0.003)	(0.006)	(0.003)	(0.005)	(0.006)	(0.003)	(0.005)
Tenure^2	-0.0006	-0.0003	-0.001	-0.001	-0.001	-0.002	-0.003	-0.001
	(0.0002)	(0.0002)	(0.0004)	(0.0002)	(0.0003)	(0.0004)	(0.0002)	(0.0004)
Female*Age							-0.030	-0.058
							(0.016)	(0.033)
Female*Age^2							0.0003	0.0009
							(0.0002)	(0.0005)
Female*Tenure	---	---	---	---	---	---	0.018	-0.003
							(0.006)	(0.008)
Female* Tenure^2	---	---	---	---	---	---	-0.001	-0.0007
							(0.0004)	(0.0006)
Selection term	-0.016	-0.077	0.032	0.164	0.439	-0.605	-0.077	0.032
	(0.036)	(0.051)	(0.094)	(0.050)	(0.197)	(0.558)	(0.050)	(0.091)
Selection term* Female	---	---	---	---	---	---	0.516	-0.637
							(0.312)	(0.583)
N.obs	10842	7933	2909	6693	3805	2888	11738	5797
N. of groups	2513	2005	835	1689	1141	810	3146	1645

Notes: All regressions also include 6 dummies for region, 2 dummies for industry and a dummy for work status (supervisor or not). Occupational controls are dropped out because in almost all cases they do not show variations through time within individuals. In the last panel, when pooled men and women are taken together, all explanatory variables are interacted with female.

Table 5c: Log Real Wage Regressions – Workers older than 40 years at the first wave (1994)
Fixed Effects Estimation with Selection Correction

	Men			**Women**			**Pooled Men and Women**	
	All	*Low*	*High*	*All*	*Low*	*High*	*Low*	*High*
Age	0.016	0.005	0.065	0.035	0.031	0.040	0.005	0.066
	(0.021)	(0.024)	(0.057)	(0.031)	(0.036)	(0.097)	(0.024)	(0.053)
Age^2	0.00008	0.0002	-0.0004	-0.0002	-0.0002	0.00009	0.0001	-0.0003
	(0.0002)	(0.0002)	(0.0005)	(0.0003)	(0.0003)	(0.0008)	(0.0002)	(0.0005)
Tenure	0.013	0.010	0.019	-0.006	0.0007	-0.038	0.010	0.019
	(0.004)	(0.005)	(0.013)	(0.006)	(0.007)	(0.016)	(0.004)	(0.012)
$Tenure^2$	-0.0005	-0.0002	-0.002	0.0003	0.0001	0.001	-0.0002	-0.001
	(0.0002)	(0.0003)	(0.0006)	(0.0003)	(0.0004)	(0.0006)	(0.0003)	(0.0006)
Female*Age							0.026	-0.025
							(0.044)	(0.123)
Female*Age^2							-.0004	0.0004
							(0.0004)	(0.001)
Female*Tenure	---	---	---	---	---	---	-0.009	-0.057
							(0.009)	(0.021)
Female* $Tenure^2$	---	---	---	---	---	---	0.0003	0.003
							(0.0005)	(0.0009)
Selection term	0.118	-0.026	-0.404	-0.055	-0.119	2.28	-0.026	-0.403
	(0.110)	(0.164)	(0.455)	(0.225)	(0.97)	(4.00)	(0.163)	(0.429)
Selection term*Female	---	---					-0.092	2.69
							(0.995)	(4.55)
N.obs	2704	2266	438	1414	1174	240	3440	678
N. of groups	592	521	104	328	300	51	821	155

Notes: All regressions also include 6 dummies for region, 14 dummies for occupation, 2 dummies for industry and a dummy for work status (supervisor or not). Workers with more than 15 years of tenure are not included since for them, the variable tenure is truncated at 15. In the last panel, when pooled men and women are taken together, all explanatory variables are interacted with female.

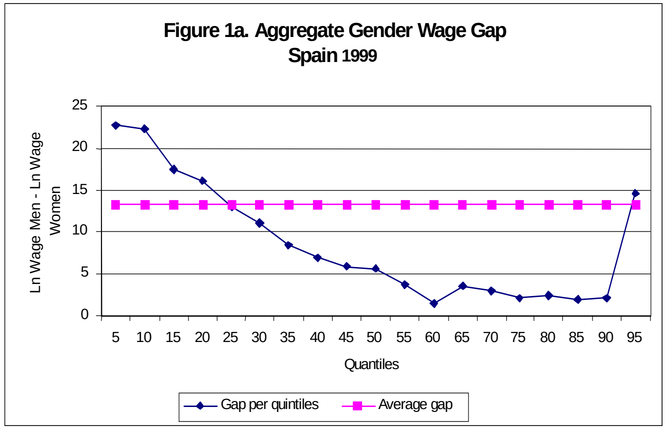

Figure 1a. Aggregate Gender Wage Gap
Spain 1999
Ln Wage Men - Ln Wage Women
Quantiles
Gap per quintiles
Average gap

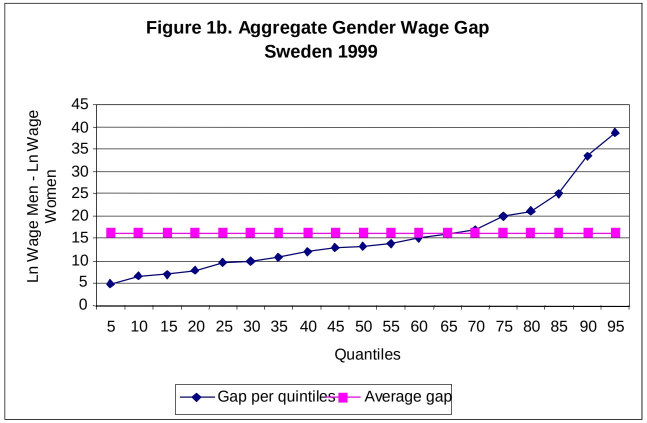

Figure 1b. Aggregate Gender Wage Gap
Sweden 1999
Ln Wage Men - Ln Wage Women
Quantiles
Gap per quintiles
Average gap

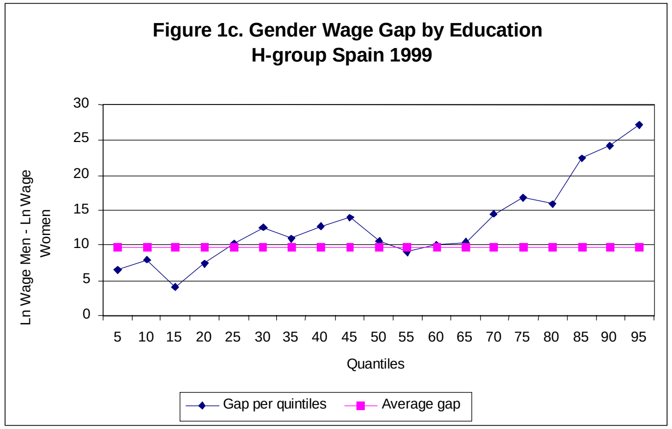

Figure 1c. Gender Wage Gap by Education
H-group Spain 1999
Ln Wage Men - Ln Wage Women
Quantiles
Gap per quintiles
Average gap

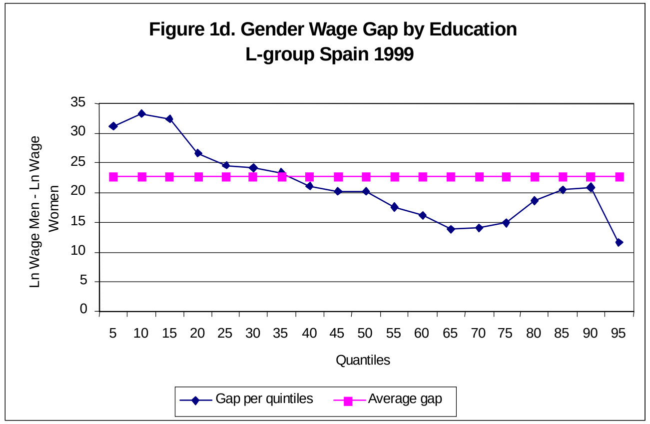

Figure 1d. Gender Wage Gap by Education
L-group Spain 1999
Ln Wage Men - Ln Wage Women
Quantiles
Gap per quintiles
Average gap

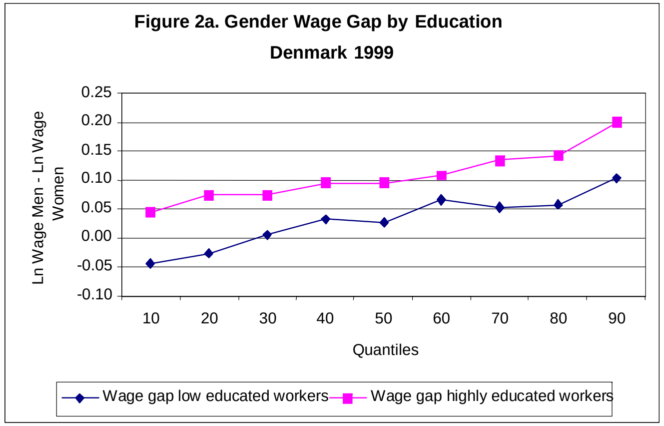
Figure 2a. Gender Wage Gap by Education
Denmark 1999
Ln Wage Men - Ln Wage Women
0.25
0.20
0.15
0.10
0.05
0.00
-0.05
-0.10
10 20 30 40 50 60 70 80 90
Quantiles
Wage gap low educated workers
Wage gap highly educated workers

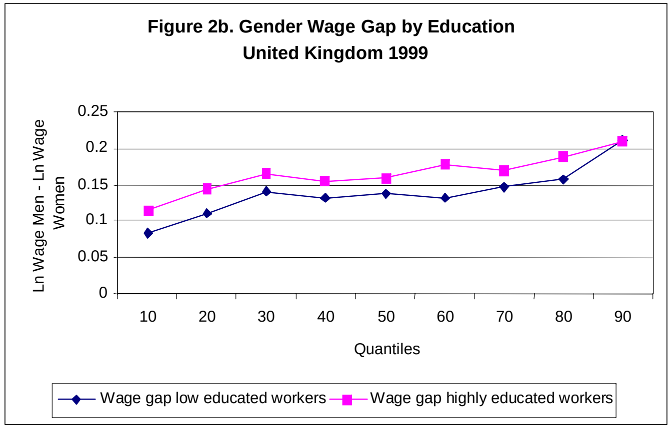
Figure 2b. Gender Wage Gap by Education
United Kingdom 1999
Ln Wage Men - Ln Wage Women
0.25
0.2
0.15
0.1
0.05
0
10 20 30 40 50 60 70 80 90
Quantiles
Wage gap low educated workers
Wage gap highly educated workers

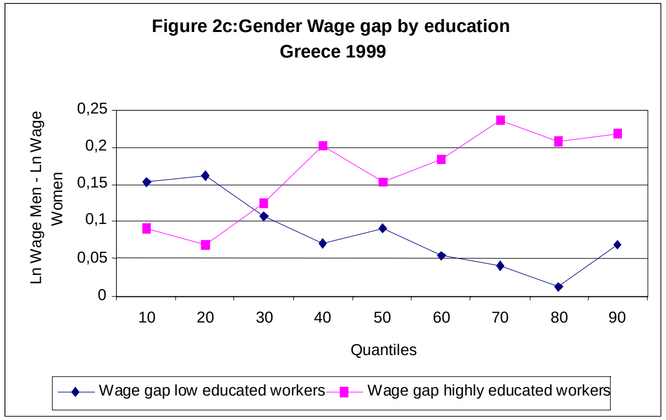
Figure 2c:Gender Wage gap by education
Greece 1999
Ln Wage Men - Ln Wage Women
0,25
0,2
0,15
0,1
0,05
0
10 20 30 40 50 60 70 80 90
Quantiles
Wage gap low educated workers
Wage gap highly educated workers

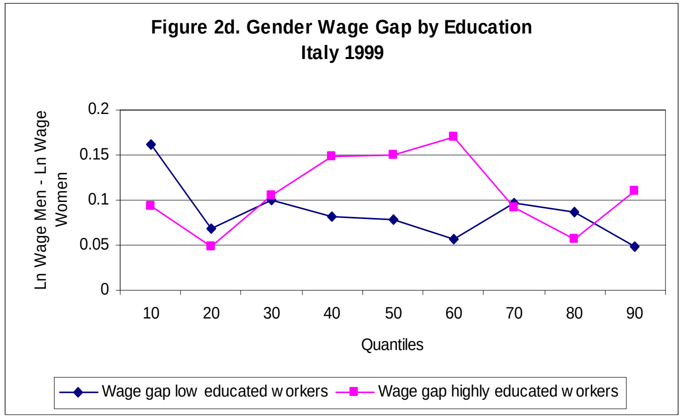
Figure 2d. Gender Wage Gap by Education
Italy 1999
Ln Wage Men - Ln Wage Women
0.2
0.15
0.1
0.05
0
10 20 30 40 50 60 70 80 90
Quantiles
Wage gap low educated w orkers
Wage gap highly educated w orkers

Figure 3. Gender gap (Observed and Counterfactual). L-group. Spain. 1999